Mitochondrial haplotype does not influence sperm motility in a UK population of men

Jim A. Mossman1,4,*, Jon Slate1, Tim R. Birkhead1, Harry D. Moore2, and Allan A. Pacey3

1Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield S10 2TN, UK 2Centre for Stem Cell Biology, Department of Biomedical Sciences, University of Sheffield, Western Bank, Sheffield S10 2TN, UK 3Academic Unit of Reproductive and Developmental Medicine, Department of Human Metabolism, School of Medicine and Biomedical Sciences, Level 4, Jessop Wing, Tree Root Walk, Sheffield S10 2SF, UK
4Present address: Department of Ecology and Evolutionary Biology, Box G, Brown University, Providence, RI 02912, USA

*Correspondence address. E-mail: jim_mossman@brown.edu

Submitted on September 28, 2011; resubmitted on November 18, 2011; accepted on November 23, 2011

Background: Sperm motility is regulated by mitochondrial enzymes that are partially encoded by mitochondrial DNA (mtDNA). MtDNA has therefore been suggested as a putative genetic marker of male fertility. However, recent studies in different populations have identified both significant and non-significant associations between mtDNA variation and sperm motility. Here, we tested whether mtDNA variation was associated with sperm motility in a large cohort of men from the UK, to test the robustness of previous studies and the reliability of mtDNA as a marker of poor sperm motility.

Methods: A total of 463 men attending for semen analysis as part of infertility investigations were recruited from a UK laboratory. Sperm motility was measured using both computer-assisted sperm analysis and traditional manual measurements. MtDNA haplogroup and haplotype were determined in 357 and 298 men, respectively, using single nucleotide polymorphism (SNP) markers throughout the mtDNA genome, and compared with sperm motility data. The linkage between the SNP markers, and possible associations between individual SNPs and motility, were also investigated.

Results: We found no statistical association between haplogroup or haplotype and sperm motility, regardless of how it was measured (P > 0.05 in all cases). Moreover, individual SNPs which were in linkage disequilibrium and dispersed across the mitochondrial genome, and therefore sensitive to mtDNA variation, were not predictive of sperm motility.

Conclusions: Mitochondrial haplotype is unlikely to be a reliable genetic marker of male factor infertility.

Key words: human / mtDNA / haplotype / infertility / sperm motility

Introduction

Approximately one in six couples are subfertile and up to half of those cases are a result of male factor infertility (Baker, 1994; WHO, 1999; Moore and Reijo-Pera, 2000). Although clinical assessment of men based on semen profiling, physical examination and hormonal and karyotype assays can often ascribe a reason for infertility, up to 50% of cases are idiopathic (Nuti and Krausz, 2008). Furthermore, up to 15% of male infertility cases may have an underlying genetic basis (Ferlin et al., 2007).

Since the mid-1990s, considerable research effort has focused on identifying genes that cause male infertility. In both human and mouse models, a number of nuclear genes linked to male infertility have been identified, particularly mutations affecting sperm function in vitro and in vivo (reviewed in Matzuk and Lamb, 2008). Moreover, a recent human genome-wide association study (Aston and Carrell, 2009) using single nucleotide polymorphisms (SNPs) identified 20 nuclear loci that were significantly associated with azoospermia and oligozoospermia. These results are promising developments for understanding the genetic basis of male infertility. However, replication of genotype association studies is essential to identify robust genetic determinants of subfertility, and only a small number of loci have been subjected to multiple-population testing, often with contradictory and/or statistically ambiguous results (Samuels et al., 2006; Pereira et al., 2007; Bandelt, 2008). This study focuses on mitochondrial DNA (mtDNA).
There is a good reason to believe that genetic polymorphisms in the mitochondrial genome may be a cause of male subfertility. First, the mitochondrial genome encodes 13 of the ~85 oxidative phosphorylation (OXPHOS) subunits essential for the production of adenosine triphosphate (ATP), which is suggested to be vital for sperm motility (Burgess et al., 2003; but see Hereng et al., 2011). Therefore, mutations in mtDNA genes encoding OXPHOS components could influence sperm motility. The second reason relates to their mode of inheritance: because mtDNA is maternally inherited, any mutations that impact on male but not female fitness cannot be effectively removed from the population by natural selection (Frank and Hurst, 1996). Therefore, mutations that negatively affect sperm motility, but have negligible effects in females, could persist in populations and even reach appreciable frequencies due to genetic drift.

There have been a number of previous investigations of mtDNA and sperm function in humans. In the context of this study, large-scale deletions (Folgero et al., 1993; Kao et al., 1995; Lestienne et al., 1997; St John et al., 2001) are not particularly relevant as they are not usually fixed within an individual or transmitted to subsequent generations, and are therefore unlikely to be a common cause of variation in sperm motility. Of greater relevance are mtDNA SNPs that have accumulated in individuals over evolutionary time and which persist in populations in the form of different mtDNA haplotypes. Two fundamentally different approaches have been taken to investigate whether mtDNA haplotypes are associated with sperm motility.

The first approach simply compares their frequency in groups of subfertile and control males, or in the absence of controls, their frequency in subfertile men with that in men from the general population (Sudjarwo et al., 2001; Pereira et al., 2005; Mohammed et al., 2006). The second, and arguably more robust approach, is to measure sperm motility in men of different mtDNA haplotype in order to compare the mean motility of the different haplotypes (Ruiz-Pesini et al., 2000; Montiel-Sosa et al., 2006; Pereira et al., 2007).

By far the best-known study of mtDNA haplotype variation on sperm motility is by Ruiz-Pesini et al. (2000), who looked in a sample of over 500 Spanish men. The authors showed that haplotype H was underrepresented and haplotype T overrepresented in men with asthenozoospermic ejaculates. Furthermore, there were significant differences in both sperm motility (as measured by ‘swim-up’) and OXPHOS component activity between different haplotypes: haplotype H was associated with more motile sperm and high OXPHOS activity, while haplotype T was associated with slower sperm and low OXPHOS activity. Initially, there was great anticipation that the results of the Ruiz-Pesini et al. (2000) study would be robust, allowing clinicians to screen for mtDNA haplotype as a predisposing factor for male infertility. In an accompanying invited editorial, Moore and Reijo-Pera (2000) noted: ‘The way that we identify and characterize the cause of untreated subfertility will be greatly impacted by the finding of Ruiz-Pesini et al. that mutations in mtDNA are associated with low sperm motility’. The Ruiz-Pesini et al. (2000) study appears to have been highly influential and to date, has been cited over 200 times.

Subsequent studies provided some support for the Ruiz-Pesini et al. (2000) findings. For example, a follow-up study in the same Spanish population suggested that haplotypes within the monophyletic haplogroup U differ in sperm motility (Montiel-Sosa et al., 2006), and two other studies have reported differences in haplotype frequency between groups of subfertile men and men with normal sperm motility (Sudjarwo et al., 2001; Mohammed et al., 2006). However, support for the role of mtDNA haplotype variation on sperm motility is by no means unequivocal. For example, Pereira et al. (2005) showed that spurious associations between mtDNA haplotype and sperm parameters can arise due to population structure: haplotype is not causative but instead reflects differences between subpopulations due to nuclear genetic or even non-genetic effects. Similarly, there are studies that have failed to identify associations between mtDNA and either sperm motility (Pereira et al., 2007) or cellular bioenergetic parameters (Amo et al., 2008). Crucially, all follow-up studies have been conducted with relatively modest sample sizes relative to the >500 men originally studied by Ruiz-Pesini et al. (2000).

Here, the results of a large study are presented with the aim of testing the hypothesis that variation in sperm motility measured by computer-assisted sperm analysis (CASA) and by routine clinical laboratory procedures (WHO, 1999), is explained by mtDNA variation. The study was conducted in a sample of almost 500 UK men undergoing diagnostic semen analysis as part of fertility investigations. This study population therefore provides an independent test of the haplotype–phenotype associations reported by Ruiz-Pesini et al. (2000), and others. This investigation simultaneously provides an opportunity to test the Frank and Hurst (1996) hypothesis which predicts that low-fitness male-specific phenotypes (e.g. poor sperm motility) can persist in populations due to the lack of opportunity for selection in males.

Materials and Methods

Patient recruitment

Semen samples were obtained from men attending the Andrology Laboratory of Sheffield Teaching Hospitals NHS Foundation Trust, UK between April and December 2006 as part of fertility investigations with their partner. The South Sheffield Research Ethics Committee approved all procedures (study reference 05/Q2305/104) and men were provided with written information about the study, in accordance with Department of Health guidelines, prior to giving informed consent. Men were excluded from the study if they: (i) had a prior history of receiving gonadotoxic therapy; (ii) were attending as part of vasectomy or vasectomy reversal procedures; (iii) had a poor understanding of English during the normal clerking process or (iv) were unable to provide an antegrade ejaculate on site by self-masturbation. Prior to sample production, all men completed a questionnaire which recorded details about their lifestyle, previous infections, use of antibiotics and the time since their last ejaculation (Supplementary data, Table S1).

Semen analysis

All ejaculates were collected into a wide-mouthed sterile container (Sarstedt Ltd., Leicester, UK). After 30 min liquefaction at 37°C, sperm parameters were evaluated according to World Health Organization (1999) to provide clinical data for the patient. Immediately following completion of the routine measurements (within 60 min of sample production), a 10 μl aliquot of neat semen and a 1:3 dilution of semen diluted in phosphate-buffered saline (PBS) were loaded into either side of a two-chambered 20 μm Microcell™ slide (Conception Technologies, San Diego, CA, USA). The sperm motility of each chamber was recorded onto videotape for later analysis by CASA. From the remaining portion of the ejaculate, a 0.5 ml aliquot of neat semen was retained and stored.
at -20 °C for later DNA extraction for mitochondrial haplotyping as outlined below.

CASA analysis

All video recordings of neat semen and 1:3 dilutions were analysed using a Hobson Tracker (Sheffield, UK) as described in Cherry et al. (2008). The image analysis settings of the CASA system were as follows: Frame rate: 50 Hz; Minimum track time: 0.8 s (40 frames); Maximum track time: 8 s (400 frames); Search radius: 6 μm; Aspect ratio: 1.49; Predict: OFF; Pause window: 0.8 s; Chamber depth: 20 μm; Analysis duration: 160 s. The output file for each sample was filtered for static and refractile debris following the methods outlined in Mossman et al. (2009), prior to statistical analysis. For each individual, three sperm populations were identified: (i) the total moving sperm population; (ii) the fastest 20% of the ejaculate judged by ranked straight line velocity (VSL), then ranked curvilinear velocity (VCL) and (iii) the fastest single sperm per ejaculate judged by ranked VSL, then ranked VCL, which represents the fastest recorded speed for that individual.

Principal component analysis

Principal component analyses (PCAs) were conducted to reduce the complexity of three collinear variables measured by CASA: (i) VCL, (ii) average path velocity (VAP) and (iii) VSL (WHO, 1999). Basically, the reason for using PCAs was to generate an index of sperm motility that was derived from the non-independent sperm motility data. Since VCL, VAP and VSL are collinear (they correlate with each other), there is redundancy in considering each variable separately, and choosing one over the others is subjective. Instead, we determined an index of sperm motility which encapsulated the majority of the motility variation in the CASA-derived variables. Three separate PCAs were conducted corresponding to the (i) total filtered sperm population, (ii) fastest 20% filtered sperm subpopulation and (iii) fastest single sperm subpopulation. The principal axis method was used to extract the principal components, and this was followed by a varimax (orthogonal) rotation. The 1st principal component (PC1), which explained the majority of the variation in VCL, VAP and VSL was used as the index of sperm motility.

DNA extraction

Each thawed aliquot of frozen semen was diluted to ~10 x 10^6 sperm/ml, based on the sperm concentration recorded at semen analysis. A 50 μl aliquot of this dilution was used for DNA extractions using a previously described plate extraction method (Whitlock et al., 2008).

Primer design

It was predicted that the haplotype variation of this population would largely reflect the 10 major European lineages (Torroni et al., 1996). Therefore, PCR primers were designed to distinguish European haplotypes as described using the Emory Group nomenclature (Torroni et al., 1996). Each haplotype alphabetical identifier represents a suite of mtDNA SNP alleles fixed in that haplotype.

Complete mtDNA sequences from previous investigations were used to identify haplotype-specific SNP loci (Andrews et al., 1999; Maca-Meyer et al., 2001; Kong et al., 2003). In total, 45 sequences (Supplementary data, Table SII) from unique and shared European and worldwide mtDNA haplotypes were aligned using Bioedit (Hall, 1999). SNP genotyping primers (Supplementary data, Table SIII) were designed following the SNP-SCALE protocol (Hinten et al., 2007). Primers were designed to type SNPs that provided the greatest resolution at discriminating known European haplotypes while also being good candidates for affecting ATP production. Putative haplotype-discriminating SNPs were identified using three processes: (i) the MITOMAP project (http://www.mitomap.org; Kogelnik et al., 1996) was consulted to screen for known haplotype-discriminating SNPs that were also non-synonymous (NS) nucleotide substitutions; (ii) haplotype-discriminating SNPs that had previously been used in the literature were used when no useful NS substitutions were present and (iii) synonymous substitutions were selected to discriminate the remaining expected haplotypes. All primers are shown in Supplementary data, Table SIII. All subsequent nucleotide numbers refer to the mtDNA revised Cambridge Reference Sequence (Andrews et al., 1999).

Polymerase chain reaction

Allele-specific PCR was conducted following the SNP-SCALE protocol (Hinten et al., 2007). All PCR reactions were performed in 10 μl volumes containing 1 X Reaction buffer (Bioline, London, UK), 0.2 mM dNTPs, 0.02 μM of each of the locked nucleic acid allele-specific oligonucleotides (Sigma-Aldrich Inc., St. Louis, MO, USA), 0.1 μM of each of the universal fluorescent oligonucleotides, 0.2 μM of the common reverse oligonucleotide (Sigma-Aldrich Inc.), 2 mM MgCl₂, 0.5 units BioTaq DNA polymerase (Bioline, London, UK) and ~10 ng of sample genomic DNA. Thermocycling was performed in a MJ Research DNA Engine Tetrad. The thermocycling profile was as follows: an initial denaturing incubation (94 °C, 3 min) followed by 10 cycles of 94 °C (30 s), 54 °C (30 s) and 72 °C (1 min). The same thermocycling conditions were repeated for another 15 cycles with an annealing temperature of 52 °C, then for a final 15 cycles with an annealing temperature of 50 °C. This was followed by a final extension step at 72 °C for 5 min. PCR products were separated on an ABI 3730 capillary sequencer (Applied Biosystems Inc., Foster City, CA, USA), and products were visualized and scored using GeneMapper Software (version 3.7, Applied Biosystems Inc.).

Assigning haplotypes and haplogroups to individuals

A total of 14 haplotype-discriminating SNPs (Supplementary data, Table SIII) were used to determine haplotype in the study samples. This was done by comparing SNP genotypes in the study population with those of individuals in the public domain for whom the entire mtDNA genome has been sequenced and haplotypes labels assigned. First, all available whole mitochondrial genome sequences of known haplotypes (n = 2826) were downloaded from the PhyloTree website (http://www.phylotree.org; van Oven and Kayser, 2009) and aligned using MAFFT version 6 (Katoh et al., 2005). After alignment, SNP genotypes at the 14 loci were extracted, and a 14-nucleotide haplotype was constructed for all of the reference sequences. The number of reference genomes per haplotype was adjusted to represent Western European rather than global frequencies reported in the literature (MITOMAP; Kogelnik et al., 1996). A comparable haplotype file with ‘sample’ genotypes at each nucleotide position was produced. Then, each sample individual was compared with the reference individuals. A haplotype was assigned to a sample individual if at least 90% of the individuals it matched in the reference data set came from a single haplotype. For example, if the 14-SNP haplotype of Individual 1 in the sample data set matched those of 10 individuals in the reference data set, all with haplotype H, then Individual 1 was scored as haplotype H. In contrast, if Individual 2 matched 10 individuals in the reference data set, of which 7 had haplotype T and 3 had haplotype U, then Individual 2 would be scored as haplotype unknown. An R script to perform this procedure is available on request.

A similar procedure was used to determine haplogroup, except haplotypes in the reference genotype file were first allocated to their respective haplogroup according to PhyloTree (van Oven and Kayser, 2009). Haplogroup compositions were as follows: L = haplotype L; M = haplotype M; D = haplotype D; N* = haplotypes I, N, W and X; R = haplotypes j
Power calculations

The power calculation to inform patient recruitment was based on the effect size estimates from the investigation by Ruiz-Pesini et al. (2000) on haplotype effects on sperm motility using a swim-up (vertical progression) test. The frequency of haplotypes H and T (the two significantly different haplotypes in their study) were different by a ratio of 9:1 and since they comprised approximately half of the total samples, the ratio of H:T other was 9:1:10. Assuming a similar haplotype structure in a UK population, a power calculation suggested that at $\alpha = 0.01$, the power of 400 patients to detect similar effect sizes was 0.91 and at $\alpha = 0.05$, it was 0.97.

MtDNA haplotype and sperm motility

Three separate analyses were conducted. First, linear mixed-effect models were used to determine the effects of haplotype and haplogroup on sperm motility, judged as PC1. Second, linear models were used to determine the effects of haplotype and haplogroup on sperm motility [PC1 and WHO A + B motility grades (WHO, 1999)]. Finally, all available PC1 and genotype data were used to examine possible associations between individual SNPs in the marker panel and motility, judged as PC1, using linear mixed-effect models.

Parametric multiple regression analyses with backward elimination methods based on significance were used to generate a minimum adequate model (MAM) using the data from semen analysis and questionnaires. Initially, all laboratory and questionnaire data (Supplementary data, Table S1) were fitted as first order terms in the model. Data were tested for normality using the Kolmogorov–Smirnov test. Sperm concentration, motile concentration and the total number of sperm (testicular volume and (iv) sperm morphology according to the WHO (1999) criteria. For PC1 sperm motility analyses, linear models were weighted by number of sperm in the analysis (‘neat’ mean $n = 528 \pm 456$ (1 SD), ‘PBS-diluted’ mean $n = 174 \pm 228$ (1 SD)) cells, since accuracy of motility estimates increases with increasing sperm numbers (Makler and Blumenfeld, 1980). Dilution was fitted as a fixed effect throughout. The terms in the MAM were: (i) dilution factor; (ii) sperm concentration; (iii) semen volume and (iv) sperm morphology according to the WHO (1999) criteria. For PC1 sperm motility analyses, linear models were weighted by number of sperm in the analysis (‘neat’ mean $n = 528 \pm 456$ (1 SD), ‘PBS-diluted’ mean $n = 174 \pm 228$ (1 SD)) cells, since accuracy of motility estimates increases with increasing sperm numbers (Makler and Blumenfeld, 1980). Dilution was fitted as a fixed effect throughout. The terms in the MAM were: (i) dilution factor; (ii) sperm concentration; (iii) semen volume and (iv) sperm morphology according to the WHO (1999) criteria.

Linkage disequilibrium between SNP alleles

Individual SNP marker associations with sperm motility were investigated (see above). It is relevant to measure the extent of linkage disequilibrium (LD) between the SNPs we typed, because if a substantial proportion of SNPs are in LD with each other, they are also likely to be informative with respect to other un-typed polymorphisms in the mtDNA genome. In other words, single SNP analyses are not only testing whether that SNP explains variation in sperm motility, but also provide information on whether other parts of the mtDNA influence motility. Two measures of LD, Hill and Robertson’s $r^2$ (Hill and Robertson, 1968) and Lewontin’s $D’$ were estimated, and a $\chi^2$-test of statistical significance, were all measured using the haploXT command (http://www.sph.umich.edu/csg/abecasis/GOLD/) implemented in GOLD (Abecasis and Cookson, 2000).

Statistical analysis

All statistical analyses and graphics were performed using the R-package software version 2.10.1 (R Development Core Team, 2009).

Results

Haplotypes and haplogroups

Of the 500 men who consented to take part in this study, 463 produced a sample that was suitable for sperm motility and genotyping analyses. All 463 semen samples were DNA extracted and at least 64.4% of the men were assigned to a known European mtDNA haplotype ($n = 298$) and haplogroup ($n = 357$) (Table I). There were 106 individuals (22.9% of DNA-extracted samples) who could not be assigned to a haplotype or haplogroup due to insufficient DNA, poor DNA quality or insufficient genotyping data. In particular, missing data at SNPs m.11467A > G and m.14766C > T made distinguishing between haplotypes T, U and V difficult.

<p>| Table I Mitochondrial haplotype and haplogroup frequencies in 463 men. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>Haplotype</th>
<th>$n$</th>
<th>$f$</th>
<th>Haplotype</th>
<th>$n$</th>
<th>$f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>236</td>
<td>0.510</td>
<td>L</td>
<td>1</td>
<td>0.002</td>
</tr>
<tr>
<td>I</td>
<td>17</td>
<td>0.037</td>
<td>M</td>
<td>1</td>
<td>0.002</td>
</tr>
<tr>
<td>J</td>
<td>25</td>
<td>0.054</td>
<td>N*</td>
<td>17</td>
<td>0.037</td>
</tr>
<tr>
<td>K</td>
<td>19</td>
<td>0.041</td>
<td>R</td>
<td>25</td>
<td>0.054</td>
</tr>
<tr>
<td>L</td>
<td>1</td>
<td>0.002</td>
<td>R0</td>
<td>292</td>
<td>0.631</td>
</tr>
<tr>
<td>M</td>
<td>1</td>
<td>0.002</td>
<td>U</td>
<td>21</td>
<td>0.045</td>
</tr>
<tr>
<td>U</td>
<td>1</td>
<td>0.002</td>
<td>Unscored</td>
<td>106</td>
<td>0.229</td>
</tr>
<tr>
<td>W</td>
<td>4</td>
<td>0.009</td>
<td>Total</td>
<td>463</td>
<td>Total</td>
</tr>
<tr>
<td>Unscored</td>
<td>165</td>
<td>0.361</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of individuals ($n$) are based on haplotypes and haplogroups scored at $>0.9$ probability, and frequencies ($f$) in each haplotype and haplogroup are based on the total number of sampled individuals.
Principal component analysis
The first principal components of the CASA variables explained 84.8, 89.4 and 87.3% of the variance in the data for the total filtered sperm population, the fastest 20% and the fastest single sperm, respectively (Table II). Comparisons with CASA variables revealed that PC1 in all populations was more highly correlated to VSL and VAP (all \( r > 0.97, P < 0.001 \)) than VCL.

Mitochondrial haplotype and sperm motility
In mixed-effect models, where the roles of ‘haplotype’ and ‘haplogroup’ were statistically assessed with terms retained in the MAM (see above) across two dilutions of semen from each male, mtDNA haplotype was not a significant predictor of sperm motility in any of the three sperm populations (all sperm, fastest 20% of sperm in the ejaculate, fastest single sperm in the ejaculate) (Table III). Similarly, mtDNA haplogroup had no effect on sperm motility judged as PC1 (\( P > 0.05 \), judged as 95% confidence intervals spanning 0). For haplotype models with repeated measures, the random effect (individual) explained 44.7, 41.4 and 37.6% of the variance in the data for the total filtered sperm population, the fastest 20% sperm subpopulation and the fastest single sperm, respectively. In other words, up to \( \sim 45\% \) of the variation in sperm motility across two dilutions of semen was explained by the inherent differences between individuals and not the terms in the model. For haplogroup models, the individual explained 46.2, 42.3 and 37.3% of the variance in PC1 data. No single haplotype demonstrated significantly very low or high sperm motility relative to the other haplotypes (Fig. 1; Table III). Because WHO A + B (WHO, 1999) motility grades were not measured in different dilutions (only neat semen), it was not appropriate to analyse these data in a linear mixed-effect model design.

The results of the mixed-effect models were confirmed by assessing the effects of haplotype in separate linear models for each of the two dilutions (neat or diluted semen). Linear models (analysis of variances) considered the role of each model term statistically, in separate dilutions. Results for neat semen are presented (Table IV) and were qualitatively the same as diluted semen (Supplementary data, Table SIV). When either haplotype or haplogroup were first order terms alone, or in combination with (i) sperm concentration; (ii) semen volume and (iii) sperm morphology according to the WHO (1999) criteria, neither haplotype nor haplogroup were significant terms in any of the models (\( P > 0.05 \) in all cases) across three subpopulations and both dilutions (Table IV, Supplementary data, Table SIV).

To conduct comparable analyses to other published studies, we also considered WHO (A + B) motility grades (WHO, 1999) as the dependent variable (sperm motility). For all models, neither haplotype nor haplogroup predicted the classification of sperm into WHO (A + B) motility grades (WHO, 1999) (Table IV and Fig. 2).

Individual SNP markers were analysed in mixed-effect models to determine if any SNP(s) were associated with sperm motility. There were no significant associations between any individual SNP genotype and sperm motility (PC1) across all sperm motility populations (Supplementary data, Table SV).

Linkage disequilibrium
Linkage analyses demonstrated substantial LD between SNPs in the marker panel. In total, 91 pair-wise comparisons were conducted and 36 of 91 marker pairs were in significant LD at \( P < 0.05 \) (Supplementary data, Table SIV). There were 60 marker pairs which had \( D^r > 0.5 \) and 4 of 91 marker pairs showed an \( r^2 > 0.3 \). All pair-wise comparisons with an \( r^2 \geq 0.013 \) (32/91 comparisons) were significant at \( P < 0.05 \). Apart from marker m.11467A > G, each SNP marker was in LD with at least one other SNP. On average, each SNP was in significant LD with 5.2 ± 2.8 (1 SD) other markers in the panel. Therefore, it is likely that if any un-typed polymorphisms in the mitochondrial genome caused variation in sperm motility, at least one of the 14 typed SNPs would be associated with both the un-typed SNP and with sperm motility.

Discussion
This investigation found no association between either: (i) an index of sperm motility (PC1); or (ii) motility measurements routinely used in the clinical laboratory (WHO A + B motility grades: WHO, 1999), and mtDNA haplotype. Even when closely related haplotypes were considered as haplogroups, no significant haplogroup effect on sperm motility was evident. Furthermore, no individual SNP typed in this investigation was associated with sperm motility. Since there was significant LD between some SNP markers, and no motility effect of any particular SNP, there is no evidence that haplotype variation, either described using the conventional nomenclature, or haplotype variation per se, has any significant effect on sperm motility. This is the first study with a comparable sample size (\( n = 357 \)) to Ruiz-Pesini

---

Table II  PCA factor loadings and eigenvalue proportions for principal components one and two in all samples (\( n = 463 \)).

<table>
<thead>
<tr>
<th>Motility parameter</th>
<th>Total sperm</th>
<th>Fastest 20%</th>
<th>Fastest sperm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PC1</td>
<td>PC2</td>
<td>PC1</td>
</tr>
<tr>
<td>Curvilinear velocity (VCL)</td>
<td>-0.583</td>
<td>-0.798</td>
<td>-0.549</td>
</tr>
<tr>
<td>Average path velocity (VAP)</td>
<td>-0.585</td>
<td>0.283</td>
<td>-0.585</td>
</tr>
<tr>
<td>Straight line velocity (VSL)</td>
<td>-0.563</td>
<td>0.532</td>
<td>-0.598</td>
</tr>
<tr>
<td>Eigenvalue</td>
<td>338.49</td>
<td>57.87</td>
<td>779.19</td>
</tr>
<tr>
<td>Proportion of variance</td>
<td>0.848</td>
<td>0.145</td>
<td>0.894</td>
</tr>
</tbody>
</table>

Three sperm populations (total sperm, fastest 20% of sperm in the ejaculate and the single fastest sperm in the ejaculate) were analysed. PBS-diluted and neat semen samples were combined for each sperm population.
et al.'s (2000) study and unlike their conclusions, no effect of mtDNA haplotype or haplogroup on any measure of sperm motility was found. These results are discussed in both clinical and evolutionary contexts. Ruiz-Pesini et al. (2000) found a significant association between mtDNA haplotype and sperm motility, judged by both WHO (1992) guidelines and a vertical progression analysis test (swim-up). They found significant associations of haplotype H and good sperm motility, and haplotype T and poor sperm motility (P = 0.031). However, their statistical analyses were post hoc and have been questioned on methodological grounds, and following further analyses using a Monte Carlo permutation test, the results were found to be non-significant (P = 0.144) (Samuels et al., 2006). The main analytical criticisms of Ruiz-Pesini et al.'s (2000) study and other mtDNA association studies [highlighted by Samuels et al. (2006)], are the non-uniform distribution of mtDNA haplotypes in populations, especially in analyses where under-represented haplotypes can lead to type I errors (Roff and Bentzen, 1989) such as when using 2 × 2 contingency tables and \( \chi^2 \) tests to address whether certain haplotypes were over-represented in men with poor (or no) sperm motility.

While the haplotype frequencies in the present study were similar to those of previously studied Western European populations (Kogelnik et al., 1996), it was not feasible in this study to determine if they were representative of haplotype frequencies in the general regional population. Therefore, it could not be determined if the frequencies of each haplotype attending clinic for fertility investigations were greater than would be expected by chance, given the overall population haplotype sub-structure.

In the present analysis, the frequency of haplotype H while high (51.0%), was similar to that observed in a previous haplotype association study with sperm motility (49.7%) (Ruiz-Pesini et al., 2000) but is possibly overrepresented when compared with haplotype frequency estimates for the UK (Cornwall: 34.8%, Mainland Britain: 35.0% and Wales: 47.8%; Simoni et al., 2000). In our sample, haplotype T [the putative poor performing haplotype in Ruiz-Pesini et al. (2000)] could not be easily distinguished from other haplotypes. However, no individual SNP was associated with high or low sperm motility phenotypes. Therefore, if unassigned haplotype T (or any other haplotype) individuals were present in our sample and significantly associated with sperm motility, it would be expected that some markers would be in LD with haplotype-specific alleles. Since there were no motility associations with any individual SNP, and significant mtDNA genome-wide LD between markers, it is unlikely that undescribed haplotypes in this sample would contribute towards high or low sperm motility phenotypes.

This is the first investigation in humans to use CASA to measure sperm motility and test for an association with mtDNA haplotype. In the context of this experiment, CASA is a more informative motility measure than the WHO (1999) grading system for several reasons.

### Table III Mixed-effect model results for haplotype and haplogroup effects on sperm motility.

<table>
<thead>
<tr>
<th>Term against model</th>
<th>Total filtered population</th>
<th>Fastest 20%</th>
<th>Fastest sperm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T (95% CI)</td>
<td>T (95% CI)</td>
<td>T (95% CI)</td>
</tr>
<tr>
<td>Neat/Dilute</td>
<td>-10.04</td>
<td>-12.63 - -8.55</td>
<td>10.34</td>
</tr>
<tr>
<td>Concentration (x10^6/ml)</td>
<td>1.13</td>
<td>-0.23 to 0.89</td>
<td>-4.95</td>
</tr>
<tr>
<td>Semen volume (ml)</td>
<td>4.98</td>
<td>1.53 to 3.59</td>
<td>-5.07</td>
</tr>
<tr>
<td>Morphology (WHO, 1999)</td>
<td>3.95</td>
<td>0.43 to 1.38</td>
<td>-3.02</td>
</tr>
<tr>
<td>Haplotype I</td>
<td>-0.11</td>
<td>-8.85 to 8.78</td>
<td>0.35</td>
</tr>
<tr>
<td>Haplotype J</td>
<td>-0.73</td>
<td>-8.10 to 3.60</td>
<td>0.15</td>
</tr>
<tr>
<td>Haplotype K</td>
<td>-0.28</td>
<td>-7.64 to 6.19</td>
<td>0.16</td>
</tr>
<tr>
<td>Haplotype L</td>
<td>0.32</td>
<td>-22.07 to 33.48</td>
<td>0.61</td>
</tr>
<tr>
<td>Haplotype M</td>
<td>-0.26</td>
<td>-32.29 to 25.13</td>
<td>0.20</td>
</tr>
<tr>
<td>Haplotype U</td>
<td>-0.54</td>
<td>-35.86 to 19.91</td>
<td>-0.04</td>
</tr>
<tr>
<td>Haplotype W</td>
<td>0.30</td>
<td>-12.29 to 15.24</td>
<td>0.10</td>
</tr>
<tr>
<td>Concentration (x10^6/ml)</td>
<td>1.41</td>
<td>-0.14 to 0.94</td>
<td>-5.42</td>
</tr>
<tr>
<td>Semen volume (ml)</td>
<td>5.04</td>
<td>1.39 to 3.08</td>
<td>-5.62</td>
</tr>
<tr>
<td>Morphology (WHO, 1999)</td>
<td>3.72</td>
<td>0.37 to 1.21</td>
<td>-2.99</td>
</tr>
<tr>
<td>Haplogroup M</td>
<td>-0.42</td>
<td>-47.90 to 31.82</td>
<td>-0.29</td>
</tr>
<tr>
<td>Haplogroup N*</td>
<td>-0.29</td>
<td>-32.82 to 26.29</td>
<td>-0.52</td>
</tr>
<tr>
<td>Haplogroup R</td>
<td>-0.46</td>
<td>-34.06 to 25.21</td>
<td>-0.58</td>
</tr>
<tr>
<td>Haplogroup R0</td>
<td>-0.36</td>
<td>-34.65 to 23.08</td>
<td>-0.59</td>
</tr>
<tr>
<td>Haplogroup U</td>
<td>-0.46</td>
<td>-36.32 to 22.70</td>
<td>-0.53</td>
</tr>
</tbody>
</table>

Models of PC1 (dependent variable) with filtered sperm population, fastest 20% population and fastest sperm are presented. For haplotype models, the random effect (individual) explained 44.7, 41.4 and 37.6% of the variance in the data for each sperm population, respectively. For haplogroup models, the random effect variance was 46.2, 42.3 and 37.3%. Confidence intervals (CIs) excluding 0 were significant at P < 0.05, judged using Markov Chain Monte Carlo sampling with 5000 iterations. Significant model terms are shown with bold Coefficients and CIs are relative to haplotype H and haplogroup L for haplotype and haplogroup models, respectively.
First, CASA tracks individual cells and measures a number of continuous cell movement characteristics known to correlate to sperm ATP levels (Irvine and Aitken, 1985; Froman and Feltmann, 1998). WHO (A+B) motility grades (WHO, 1999) simply classify the proportion of an ejaculate that swims with VCL ≥ 5 μm/s, and therefore is not a direct measure of sperm motility characteristics. Second, in motility analyses, where static cells are filtered from the sperm population (Mortimer and Mortimer, 1988; Holt et al., 1996; Mossman et al., 2009), CASA measures only cells that are moving. However, because static (dead or quiescent) sperm may be a biologically significant determinant of (in)fertility (Larsen et al., 2000), here, we separately investigated two measures of sperm performance: (i) WHO A+B motility grades (WHO, 1999); and (ii) an alternative continuous measure of motility, PC1.

The significant terms retained in our analysis model reflected biologically important semen parameters for sperm motility and fertility, lending support for their use in a clinical context (and vice versa) (WHO, 1999; Cooper et al., 2010). Since previous haplotype association studies were also conducted in clinical laboratories providing additional semen analysis data, it is perhaps surprising that other groups did not report them or use them in their analyses.

Sperm phenotype is inevitably the product of genetic and environmental factors, and possibly genetic × environmental interactions. Large within-haplotype variation in PC1 in this study suggests there could be nuclear genetic variation responsible for sperm motility differences as nuclear background could not be controlled in this experiment. Alternatively, as suggested by Montiel-Sosa et al. (2006), there may even be variation within a haplogroup that provides functional variation in sperm motility (i.e. sub-lineages within the same haplogroup have significantly different sperm phenotypes). In the present study, mtDNAs were screened at a minimal resolution to allow European haplotypes to be delineated. However, haplogroups are families of molecules that are closely related according to their position in a phylogeny. Nested within haplogroups, are more specific haplotype variants, and these are only identified by high-resolution screening (e.g. genotyping at a larger number of markers or sequencing at highly polymorphic loci). In this study, it was not feasible to genotype or sequence all individuals to a very fine-scale. We are therefore unable to verify the findings of Montiel-Sosa et al. (2006) that significant variation in sperm motility occurs between haplotypes, within haplogroups, as a result of sequence variation.

It is possible for spurious haplotype–phenotype associations to arise if there is underlying genetic structure in the population (Pereira et al., 2005) (i.e. by concealing the exact role of mtDNA variation because of potential LD with nuclear DNA variants). In the study of Ruiz-Pesini et al. (2000), the positive association between haplotype and sperm motility may reflect an inter-population variation in motility, since nuclear genes or even a common environmental background could not be controlled for. That Ruiz-Pesini et al. (2000) found a positive association between haplotype and sperm motility yet the present study did not, may suggest that there are different degrees of population structure between the two samples. The discordance between the present study and that of Ruiz-Pesini et al. (2000) is consistent with other studies which suggest ethnicity (Helgason et al., 2006), and...
Table IV  Linear models of haplotype and haplogroup effects on neat sperm motility.

<table>
<thead>
<tr>
<th>Model</th>
<th>Term against model</th>
<th>d.f.</th>
<th>Sum-of-squares</th>
<th>F</th>
<th>P-value</th>
<th>Sum-of-squares</th>
<th>F</th>
<th>P-value</th>
<th>Sum-of-squares</th>
<th>F</th>
<th>P-value</th>
<th>WHO A + B %d</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Concentration (× 10^6/ml)</td>
<td>1</td>
<td>1 281 278</td>
<td>27</td>
<td>&lt;0.001</td>
<td>2 428 665</td>
<td>22</td>
<td>&lt;0.001</td>
<td>2660</td>
<td>0.01</td>
<td>0.91</td>
<td>2505</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 797 369</td>
<td>7</td>
<td>&lt;0.01</td>
<td>1 075</td>
<td>8</td>
<td>&lt;0.001</td>
<td>3085</td>
<td>3</td>
<td>0.07</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>808 269</td>
<td>3</td>
<td></td>
<td>799</td>
<td>1</td>
<td>0.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2505</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Concentration (× 10^6/ml)</td>
<td>1</td>
<td>1 328 847</td>
<td>27</td>
<td>&lt;0.001</td>
<td>2 353 878</td>
<td>21</td>
<td>&lt;0.001</td>
<td>1 8982</td>
<td>0.08</td>
<td>0.78</td>
<td>3185</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 794 092</td>
<td>8</td>
<td>&lt;0.001</td>
<td>2 2229</td>
<td>17</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 8982</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Haplotype (WHO, 1999)</td>
<td>1</td>
<td>538 850</td>
<td>31</td>
<td>&lt;0.001</td>
<td>1 385 436</td>
<td>17</td>
<td>&lt;0.001</td>
<td>933 476</td>
<td>4</td>
<td>&lt;0.05</td>
<td>2124</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 974 092</td>
<td>8</td>
<td></td>
<td>2229</td>
<td>17</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>933 476</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Haplogroup</td>
<td>1</td>
<td>260 074</td>
<td>1</td>
<td>0.39</td>
<td>623 305</td>
<td>1</td>
<td>0.34</td>
<td>997 246</td>
<td>1</td>
<td>0.51</td>
<td>707</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 262 777</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 262 777</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Motility was judged as PC1 in the total filtered sperm population, fastest 20% population, fastest sperm population and WHO (A + B) motility grades. Four models are presented: MAMs (Models 1 and 2) and haplotype and haplogroup alone (Models 3 and 4, respectively). Haplotype or haplogroup were non-significant predictors of motility measurements in all models and across all motility variables. Values in bold are significant at P < 0.05. Differences in degrees-of-freedom between models reflect unobservable model terms values or PC1 data in some samples.

Model 1: d.f. (10 272) except WHO A + B analysis (10 281); r^2 = 0.23, F = 7.90, r^2 = 0.20, F = 6.98, r^2 = 0.06, F = 1.83, r^2 = 0.23, F = 8.19.
Model 2: d.f. (8342); r^2 = 0.21, F = 10.85, r^2 = 0.19, F = 9.64, r^2 = 0.05, F = 2.32, r^2 = 0.19, F = 10.13.
Model 3: d.f. (7281); r^2 = 0.03, F = 1.11, r^2 = 0.03, F = 1.27, r^2 = 0.02, F = 1.00, r^2 = 0.03, F = 1.09.
Model 4: d.f. (5340); r^2 = 0.02, F = 1.32, r^2 = 0.02, F = 1.49, r^2 = 0.02, F = 1.11, r^2 = 0.02, F = 1.24.
variable effect sizes in different study populations (Samuels et al., 2006), can affect the repeatability of genotype–phenotype associations.

To date, all investigations on the role of mtDNA haplotype on sperm motility have been conducted on men attending fertility clinics. It is therefore unclear whether the results of these studies are representative of the general population. Alternative conclusions from different studies suggest however, that population structure probably has a greater influence on results than the demographic of men being sampled, since a similar sample of possibly subfertile men were studied in all cases. Matching cases and controls (e.g. Pereira et al., 2005, 2007). Possible reasons for discordance between haplotype association studies have been highlighted, and probably reflect the difficulty in matching cases and controls, and existence of genuine variation in genetic effect sizes between populations (Samuels et al., 2006). In human sperm studies, these potentially confounding factors are difficult to avoid, but there is a growing body of evidence in ideal laboratory systems that suggest variation of cellular bioenergetics (Carelli et al., 2002; Moreno-Loshuertos et al., 2006; Amo et al., 2008) and sperm performance (Dowling et al., 2007; Friberg and Dowling, 2008; Mossman et al., 2010) are independent of mtDNA haplotype. As a cautionary note, the repeatability of haplotype association studies may be poor, but at the same time, poor repeatability per se is a useful result when deciding which genetic tests should not be used in the clinical environment as diagnostic tools for male factor infertility.

Figure 2 Boxplot of mitochondrial haplogroup effects on human sperm motility [WHO A + B motility grades (%)]. Results were non-significant when samples were considered as individual haplotypes or grouped as haplogroups (P > 0.05, shown, see Table IV).
Supplementary data
Supplementary data are available at http://humrep.oxfordjournals.org/.

Acknowledgements
The authors thank D. Smith, M. Fletcher, D. Saxton, D. Taylor and T. Hart for their immeasurable help in the Andrology Laboratory. Geoff Holson provided support for CASA set-up. E.Gillies provided useful comments on the manuscript. The authors thank two anonymous reviewers for comments which greatly improved the manuscript.

Authors’ roles

Funding
J.A.M. was supported by a University of Sheffield PhD studentship; J.S. was funded in part by the BBSRC. T.R.B. was supported by a grant from the Leverhulme Trust.

Conflict of interest
None declared.

References
Mitochondrial haplotype and sperm motility


St John JC, Jokhi RP, Barratt CLR. Men with oligoasthenoteratoozoospermia harbour higher numbers of multiple mitochondrial DNA deletions in their spermatoozoa, but individual deletions are not indicative of overall aetiology. Mol Hum Reprod 2001; 7:103–111.


